CERTAIN PROPERTIES OF INDEX NUMBERS*

By Truman L. Kelley, *Stanford University*

Acting upon the assumption that the significance, limitations, and general meaning of an index number lie within the grasp of a non-economist, the present writer has dared to enter the field, already so ably investigated by economists, with the present statistical contribution. The two ideas that he brings to this study, which have been but partially noted in the references given at the close of the article, are, first, the importance of knowing the probable error of index numbers and the method of their calculation; and, second, the significance of correlation between price ratios in the building up of an index and in the entry and withdrawal of items.

The price of a commodity in some one year, $p^1{}_1$ (the superscript designates the commodity, while the subscript designates the year), divided by the price of the same commodity in a second year, $p^1{}_2$, is $p^1{}_1/p^1{}_2$, and is called a price ratio. A composite of several such ratios purporting to portray a general relationship between prices in the two years is a price index, P_1/P_2. The fundamental concept in this is the ratio or geometric concept. Indices can be built upon many bases, but irrespective of the method of construction, the usual interpretation will involve this geometric concept. The lay reader will think that P_1 is a certain proportion of P_2, and P_2 is the inverse proportion of P_1. An index which is not reversible does not parallel the thought processes inherent in the concept "price ratio," and this more elementary concept, where reversibility is the rule, is the one by means of which "price index" is interpreted. Even writers who are quite aware that the index they are using is not reversible, use price ratios and price indices in such a way that it is obvious they expect the same sort of concept to be called up in the reader's mind; for example, "$p^1{}_1/p^1{}_2=122$, but $P_1/P_2=120$ so that, etc."

In as far as the concept P_1/P_2 is commonly of a different nature from $p^1{}_1/p^1{}_2$, it lies in the fact that P_1 and P_2 are averages, and $p^1{}_1$ and $p^1{}_2$ are single measures. Accordingly, to parallel customary think-

* The article herewith presented, except for certain footnotes and a few minor changes, was written before the appearance of the March, 1921, number of the Quarterly containing the abstract of Professor Fisher's Atlantic City paper. This paper and the attendant discussion have suggested certain things which might well have been included, but they have not indicated a withdrawal of any portion of the present paper. It therefore seems more convincing, as an independent contribution, to let the article stand as originally prepared than to rewrite it with a view to bearing upon the detailed questions raised by Professor Fisher.

ing, P_1/P_2 should mean a reversible proportion between averages. What an "average" is may not be so definitely established in the minds of scientific people generally as is the idea "ratio," but probably the most common concept is that of arithmetic average or mean. We therefore have the somewhat anomolous situation of P_1/P_2 calling up the arithmetic concept when dealing with the two separate elements involved in it, but the geometric concept when dealing with the thing entire. Since this mixture of concepts seems likely to persist, the writer proposes as an important test of the excellence of an index number the closeness with which the operations involved in it parallel general thinking tendencies: *First and most important, reversibility of ratio, and second, arithmetic averages involved in the parts.*

That a price index has a probable error is a fact not always recognized and not entirely obvious, for it may easily happen that the price ratios are entirely reliable. It may be possible to say that the price of cotton at a certain time was $p^1{}_1$ and at a second time $p^1{}_2$. If the price quotations are accurate, then the price ratio $p^1{}_1/p^1{}_2$ is a true measure. The average of several such gives P_1/P_2, which is invariable. Therefore, P_1/P_2 has zero probable error as far as being the average of these particular things, but the very combining of them involves the assumption that the index has significance beyond the particular data from which it is calculated. The only exception would be when P_1 and P_2 are determined from all the possible data. As an example, let $p^1{}_1$ be the price of coal at a certain mine at the first date, $p^2{}_1$ the price at a second mine, . . . , $p^n{}_1$ the price at the last mine, and similarly for the p_2's. Then, since all the sources are involved, P_1/P_2 is the index of coal prices and has no probable error, except such as might be due to faulty quotations and calculations and could therefore, by proper care, be made negligible.

This is not the typical situation. Ordinarily but a few quotations are worked up into an index and the result taken as representative of an industry or a field. We therefore have quotations which are samplings of the prices in the industry, and the statistical methods for determining the reliability of samplings apply. The formulae for probable errors given later in this article are based upon certain assumptions, including that of random sampling, but if 25 or more per cent of the possible quotations are utilized, material error in the formulae is introduced, the true probable errors being less than those given by the formulae. It is to be understood that by probable error in an index number is meant that which arises from incompleteness of data. In the following determinations of probable errors of index numbers as given by various formulae, the attempt is to see how closely one can

approximate, by a sample, the number which would be obtained were all the possible data utilized in determining the same sort of index. The probable error indicates how closely the results from the sample may be expected to tally with the results from the whole. Should there be a constant tendency in the form of index used, systematically leading to too high or too low a value, we have a systematic error, which is entirely distinct and which is not measured by the size of the probable error.*

The reason why a few quotations can yield an index which is a close approximation to a general tendency is that there is a high correlation between the quotations included and those not included in the index but pertinent to the function being measured. If there are two hundred coal mines and quotations from a half dozen are taken, an index in close agreement with the true index based upon the two hundred may be expected, because of the high correlation between quotations at different mines. To say that there is a high correlation is not equivalent to saying that the prices at the different mines tend to approach the same level, but that they tend to maintain a uniform difference. Mine A, near tidewater, may sell at a certain price, p^1, much higher than that, p^2, at mine B, remote from a center of consumption, without indicating an economically abnormal condition in the coal trade. If p^1, p^2, and other similar measures are averaged, the probable error of this average is not given by the usual formula

$$\text{P. E.}_{\text{mean}} = .6745 \frac{\sigma}{\sqrt{N}}$$

due to the heterogeneity of, and to the correlation between, the p's. As an illustration, more extreme than mine quotations on coal, let us average the following prices:

Bacon per pound	$.70
Bread per pound	.10
Potatoes per bushel	1.20
Apples per box	10.00
Average	$3.00
Standard deviation	4.06
P. E. (by above formula)	1.37

* In the tests of indices suggested later in the paper there will be found none to the effect that an index should have no bias. The reason for this is that reversibility of ratio, or change of base, which is included as one of the tests, is not possible with a "biased" index. Fisher (1921) shows that an index may possess a bias due to form and a second bias due to base value weighting, and that these may exactly neutralize each other. Such a situation would, statistically, be the same as one not involving bias.

Now, presumably, the probable error of no single one of these quotations is as great as \$1.37, and the average of them all will probably fluctuate but little. There probably is positive correlation between these food prices, a rise in one generally going with a rise in each of the others. These conditions are not those under which the probable error of an average is given by the usual formula. For statistical purposes there is much to be gained by having homogeneous uncorrelated material. We can secure measures which are nearly, if not entirely, homogeneous and uncorrelated by dealing with price ratios instead of prices.*

Accordingly, if the price index showing prices in year 1 relative to year 2, called i_{12}, is given by the equation,

$$i_{12}=\frac{P_1}{P_2}=\frac{1}{N}\Sigma\frac{p_1}{p_2} \quad \text{(Index formula 1)}$$

and if the standard deviation of the price ratios is σ_{12}, the probable error of i_{12} is given by

$$P.\ E.\ i_{12}=.6745\frac{\sigma_{12}}{\sqrt{N}}. \quad \text{(Probable error of index formula 1)}$$

Let us consider another kind of index,

$$i_{12}=\frac{P_1}{P_2}=\frac{\Sigma p_1}{\Sigma p_2}. \quad \text{(Index formula 2)}$$

The complete probable error formula for this kind of index involves the correlation between the p's. (See Pearson, 1910.) The index,

$$i_{12}=\frac{1}{\Sigma_w}\Sigma\left(\frac{p_1}{p_2}\right)w \quad \text{(Index formula 3)}$$

will be more reliable than formula 1 if the weights, w, used are exactly or approximately proportionate to the values of the commodities involved. In general, the greater the price ratio the less the consumption and vice versa, so that the distribution of the weighted price ratios will

* In one sense, both prices and price ratios are very highly correlated, but these correlations have quite different statistical consequences. As the price of coal at mine A approaches p^1_1, due to correlation the price at mine B approaches what may be a very different value, p^2_1; but as the ratio, p'_1/p^1_2, from the quotations of mine A approaches, as time changes, the value ρ, due to correlation, the ratio of the quotations from mine B may be expected to tend toward the same value ρ. (The rigorous proof of this statement would be necessary before the present treatment and statement of probable errors can be considered final. Whatever error is involved is of a conservative nature, as it almost certainly would tend to make the obtained probable errors too large.) Although correlation between prices tends to throw ratios together, it tends to keep prices apart. If, therefore, we deal with ratios, the effect of correlation has already operated upon the measures used, making the distribution of ratios more homogeneous, and as a consequence making the mean more reliable. In other words, the standard deviation of the ratios of prices at date 1 to those at date 2, σ_{12}, is reduced from what it would be were there no correlation between prices, so that by this very reduction, the probable error formula when applied to ratios takes account of the correlation between prices at two different dates. For a rigorous approach to the question of probable error of a ratio see Pearson (1910 and 1911).

have a smaller variability than the distribution of price ratios alone. If $w = p_2q_2$, the value of the transactions in year 2, the formula becomes

$$i_{12} = \frac{\Sigma p_1q_2}{\Sigma p_2q_2}. \quad \text{(Index formula 4)}$$

Formula 4 is but a type of formula 3. It is undoubtedly more reliable than either 1 or 2, but there are too many variables involved for the writer to attempt a calculation of its probable error based upon the data for two dates only. If, however, the commodities are divided into random halves and indices determined from each half, the correlation between these sub-indices may be calculated, and from it the probable error of the total index may be obtained, as follows:

Let there be n commodities, equally excellent as representative of the whole field, which are built up into the index i. In order to determine the probable error of i we may first build up two indices, A and B, each based upon a random half of the commodities. Calculation of A and B for a number of dates will give two series, the correlation between which may be found. In doing this it is desirable that the time interval between successive indices be sufficient to insure the relative independence of the commodity quotations involved. Just as the average of the prices of bread on January 1 of a certain year and on December 30 of the same year will in general give a truer average yearly price than the average of the prices on June 30 and July 1, because in the former case the two quotations are nearly independent while in the latter one has practically but a single quotation, so sub-indices calculated at too short intervals of time scarcely constitute new data, but rather repetitions of old data. Were the correlation between successive quotations known, practical limits could be set giving periods shorter than which it would not be worth while to calculate sub-indices. Having r, the Pearson product-moment coefficient of correlation, between these sub-indices, ordinarily called the reliability coefficient, we may infer the reliability coefficient, R, of the entire index, i. This is a measure of the extent to which the index i would be expected to correlate with a second similarly derived index, and is given by the formula (Brown, p. 102, 1911),

$$R = \frac{2r}{1+r}. \quad \text{(Formula to infer the reliability of a measure knowing the reliability of its halves)}$$

The probable error of i is then given by*

$$P.E._{\cdot i} = .6745\,\sigma'\sqrt{1-R}$$

* The correlation between the index i and the "true" index, where the true index is defined as the average of a very large number (an infinite number) of such indices as i, is $\sqrt{R}$. (For proof see Kelley, 1916.) Now consider a correlation table, or scatter diagram, between true index values and i-values. The standard deviation of the arrays corresponding to a certain true score, according to the usual

in which σ' is the standard deviation of the indices i for the same period of time as covered by the correlation table giving r. If σ equals the average of the standard deviations of the two series of sub-indices (presumably these two standard deviations are very nearly equal) then*

$$\sigma' = \sigma \sqrt{\frac{1+r}{2}}.$$

Substituting the values found for R and σ', the probable error formula may be rewritten,

$$P.\ E._{\cdot i} = .6745\ \sigma \sqrt{\frac{1-r}{2}}.$$ (Probable error of an index in terms of the correlation and standard deviation of sub-indices)

Note that r and σ must be obtained from the same series of sub-indices.

The practical advantages of reporting two sub-indices as well as the total index may well be as great as has been found to be the case in reporting two comparable measures in the fields of psychology and education. The probable error of any index may be determined if comparable sub-indices are calculated and if the series of indices covers a sufficient length of time to yield a reliable measure of correlation between sub-indices. Probably 16 pairs of quarterly sub-indices would suffice.

Accordingly, a second important measure of the excellence of an index number is *the size of its probable error.*†

---

formula for the standard deviation of an array (see Yule, 1912) is, $\sigma_{i.t} = \sigma'\sqrt{1-(\sqrt{R})^2}$, in which σ' is as defined above and $\sigma_{i.t}$ is the standard deviation of the i's for a given true value of the index. Thus $\sigma_{i.t}$ is simply the standard error of the index i, and the probable error is .6745 times as great.

* Let σ_1 = the standard deviation of the A series of indices, σ_2 of the B series, and let $i=(A+B)/2$. Let Δ stand for a deviation (an error as judged by the true index) in the i index, δ for one in the A index, and d for one in the B index, then,

$$2\,i = A+B$$
$$2\,\Delta = \delta + d$$

Squaring, summing, dividing by N, the number of cases, and noting that the sum of the δd products equals $Nr\sigma_1\sigma_2$, yields,

$$4(\sigma')^2 = \sigma^2_1 + \sigma^2_2 + 2r\sigma_1\sigma_2.$$

If the standard deviations in the right-hand member are nearly equal we may replace them by $\sigma[=(\sigma_1+\sigma_2)/2]$ giving,

$$\sigma' = \sigma\sqrt{\frac{1+r}{2}}.$$

† I judge from the none too complete abstract that Fisher (1921) has calculated a large number of different indices from the same material and found that certain formulae give highly comparable results. The uniformity of indices involving the same data is not the problem of reliability here attacked. We are concerned with the problem of sampling. As to whether Professor Fisher has also compared an index determined from a part of his data with the same index as obtained from a larger part I cannot determine from the abstract, but if so it constitutes an experimental approach to the problem in hand. One would expect that the differences which Professor Fisher would find between an index based upon, let us say, ¼ of his data and one based upon the remaining ¾ would be somewhat larger than implied by the formulae here given, as the index based upon the ¾ would be a fallible standard. A study of the uniformity of indices based upon the same data throws light upon the existence and the nature of systematic tendencies, or biases, but none whatever upon the error of sampling.

Space will not permit a discussion of the probable errors of all the proposed types of indices, but to point out the necessity of such discussions the writer has made an estimate, after more or less complete mathematical analysis, of the relative size of the probable errors of the index numbers given in the table on pages 836–7.

The one that seems the most reliable of all, and that also most completely meets other conditions except that of paralleling general thinking tendencies, is the weighted geometric mean index, in which the weights are roughly proportional to the reliabilities of the price ratios. This requirement as to weights is practically no limitation at all, as it is regularly approximated to by customary weighting devices. Practically without exception the observations of Mitchell (1915) as to what items to include in an index and what weights to give, are statistically equivalent to weighting price ratios according to reliability. As soon as a commodity becomes archaic the proper thing to do is to withdraw it, and withdrawals and entrances are readily accomplished with the geometric index. The weighted geometric mean index formula is

$$i = \sqrt[\Sigma w]{\frac{(p^1{}_1)^{w_1}\,(p^2{}_1)^{w_2}\ \ldots\ (p^n{}_1)^{w_n}}{(p^1{}_2)^{w_1}\,(p^2{}_2)^{w_2}\ \ldots\ (p^n{}_2)^{w_n}}}. \quad \text{(Index formula 5)}$$

For convenience, and without any loss of generality, Σw may be made to equal 1. Thus, letting $\omega_1 = w_1/\Sigma w$, $\omega_2 = w_2/\Sigma w$, etc., and as before. $\rho_1 = p^1{}_1/p^1{}_2$, etc.,

$$i = \rho_1^{\omega_1}\,\rho_2^{\omega_2}\ \ldots\ \rho_n^{\omega_n}. \quad \text{(Index formula 5a)}$$

Note that with this formula the index is reversible and that there is complete freedom in changing the base. Assuming as before that there is no correlation between ratios, the probable error is given by

$$P.\,E._i = .6745\frac{i}{\Sigma w}\sqrt{\frac{w^2{}_1\sigma^2{}_1}{\rho^2{}_1} + \frac{w^2{}_2\sigma^2{}_2}{\rho^2{}_2} + \ \ldots\ \frac{w^2{}_n\sigma^2{}_n}{\rho^2{}_n}} \quad \text{(Probable error of the weighted geometric mean index)}$$

in which the ρ's are successive price ratios and the σ's their standard deviations. As an approximation, the σ's may be considered to be equal to each other and to equal the standard deviation of the distribution of price ratios. In order that this probable error remain small, it is necessary that no one of the ratios w_1/ρ_1, w_2/ρ_2, etc., be exceptionally large.

$$\frac{w_1}{\rho_1} = \frac{w_1 p^1{}_2}{p^1{}_1}.$$

Letting $q^1{}_1$ equal the quantity of the commodity consumed, or in trade, it would be expected that $q^1{}_1p^1{}_1$ would fluctuate much less than $p^1{}_1$, and whereas there might be danger of $p^1{}_1$ becoming extremely small or large there is not equal liklihood of $q^1{}_1p^1{}_1$ doing so. Accordingly, if w_1 is approximately$=q^1{}_1p^1{}_1$, then $w_1/\rho_1=q^1{}_1p^1{}_2$, a magnitude which is not likely to be extremely large. However, should a commodity change greatly in its relative importance, the weighting of it may easily be changed as follows:

Let it be desired to change the weight of the price ratio ρ_1 from w_1 to W_1, which we will say is a smaller weight. We need not impose the condition that $\rho_1=i$. For $\rho_1 > i$ we will search the list of price ratios for (a) a ratio $> i$ which is underweighted, or (b) a ratio $< i$ which is overweighted. Suppose ρ_2 is such a ratio. Ordinarily there are a number of price ratios$=1.0$, or i, or some other value which is the modal value. These may be combined and represented by ρ^s, where ρ is this modal value and s the sum of the weights of all the ratios having this value ρ. Letting P stand for the product of all the terms other than ρ_1, ρ_2, and the ρ terms, we have

$$i=\sqrt[\Sigma w]{\rho_1{}^{w_1}\,\rho_2{}^{w_2}\,\rho^s\,P}$$

and it is desired to change this to

$$i=\sqrt[\Sigma W]{\rho_1{}^{W_1}\,\rho_2{}^{W_2}\,\rho^S\,P}.$$

The first index will equal the second in case

$$w_1+w_2+s=W_1+W_2+S \tag{1}$$

and also,

$$\rho_1{}^{w_1}\,\rho_2{}^{w_2}\,\rho^s=\rho_1{}^{W_1}\,\rho_2{}^{W_2}\,\rho^S, \tag{2}$$

or, taking logarithms,

$$w_1\log\rho_1+w_2\log\rho_2+s\log\rho=W_1\log\rho_1+W_2\log\rho_2+S\log\rho.$$

W_1 is the new weight that has been assigned (this may be zero) so that everything involved is known except W_2 and S, and the solution of the two equations simultaneously will yield these. Ordinarily S will differ but slightly from s, and W_2 will differ from w_2 in the direction in which it is desirable it should differ. Thus, as a practical matter, the weight of any price ratio, whether equal to i or not, may be changed without affecting the index.

No other index, as far as the writer can determine, offers the extreme flexibility in changing weights, dropping or adding new items, here found to exist for the geometric mean index. Since this is so, the

weights can be made such that extreme ratios are given small weights or eliminated. As a consequence, the probable error of such a weighted geometric mean index may be expected to be smaller than that of any other index mentioned. The excellence of this index seems to the writer so great as to warrant its use, even though it involves a change in the established habits of interpretation of the usual reader.

Two criteria, the paralleling of habitual modes of thinking and reliability, have been proposed in judging the excellence of an index measure. Fisher (1913) has used eight other tests, three of them being tests only of "trade" indices. It would seem that these latter would be of particular importance only in case an index ceases to be a sampling and becomes an expression of the sum total of transactions involved. The table on page 836–7, in part taken from Fisher (1913), gives "scores" of the most important index measures upon several tests or criteria of excellence.

Test 1: Reliability. In giving scores upon this point the writer has freely used his judgment in the case of indices for which no simple probable error formula is available. More or less complete statistical analysis has preceded this scoring, but it is in no sense to be considered final. An "*s–i*" after a score means that no simpler way for calculating the probable error than by means of the correlation between comparable sub-indices seems to be available. As the writer judges this test to be the most important of all, the scoring is 3, 2, 1, and 0, instead of 2, 1, and 0—the larger the score, the higher the rating.

Test 2: Parallels habitual modes of thinking. Score 2, 1, 0.

The following tests are from Fisher.

Test 3: Proportionality. "A price index should agree with the price ratios if these all agree with each other." Stated algebraically:

$$\text{Given } \frac{p^1{}_1}{p^1{}_2} = \frac{p^2{}_1}{p^2{}_2} = \text{etc.} = i. \quad \text{Required that } \frac{P_1}{P_2} = i.$$

Score of 2 if true for any two years. Score of 1 if true only when year 2 is the base year.

Test 4: Entry and withdrawal. A price index should permit the entry and withdrawal of price ratios without changing the value of the index. Fisher uses a less general test: "A price index should be unaffected by the withdrawal or entry of a price ratio agreeing with the index." The scoring here follows Fisher, except for formula 5,

which Fisher does not include in his list of 44, and for formulae 14 and 15 which are here scored higher than by Fisher.* Score 3, 2, 1, 0.

* Fisher scores both of these formulae zero on the basis of entrance and withdrawal of items. It is, however, a simple matter to enter or withdraw price ratios, if the proper choice of weights is made. Proof for formula 15 is as follows:

To simplify notation, let

$$a=\Sigma p_1q_1$$
$$b=\Sigma p_2q_1$$
$$c=\Sigma p_1q_2$$
$$d=\Sigma p_2q_2$$

Then,

$$i=\sqrt{\frac{\Sigma p_1q_1}{\Sigma p_2q_1}\times\frac{\Sigma p_1q_2}{\Sigma p_2q_2}}=\sqrt{\frac{ac}{bd}}. \quad \text{(Index formula 15)}$$

Consider first the case of entering a price ratio which agrees with the index,

$$\rho_1=\frac{p^1{}_1}{p^1{}_2}=i,$$

and let it be desired to enter it with the weight q_1 in year 1. It is required to determine q_2 so that this new price may be included without changing the value of the index. Since $\frac{p^1{}_1}{p^1{}_2}=\frac{p^1{}_1q^1{}_1}{p^1{}_2q^1{}_1}=\frac{\sqrt{ac}}{\sqrt{bd}}$ therefore $p^1{}_1q^1{}_1$ is equal to some constant, k, times $\sqrt{ac}$, and $p^1{}_2q^1{}_1$ is equal to the same constant times $\sqrt{bd}$, so we may write,

$$\left.\begin{array}{ll} p^1{}_1q^1{}_1=k\sqrt{ac}, & p^1{}_1q^1{}_2=j\sqrt{ac} \\ p^1{}_2q^1{}_1=k\sqrt{bd}, & p^1{}_2q^1{}_2=j\sqrt{bd} \end{array}\right\} \quad \text{from which } \frac{q^1{}_1}{q^1{}_2}=\frac{k}{j}.$$

Introducing the new price, we have

$$\sqrt{\frac{a+k\sqrt{ac}}{b+k\sqrt{bd}}\times\frac{c+j\sqrt{ac}}{d+j\sqrt{bd}}}=\sqrt{\frac{ac}{bd}}.$$

It remains only to solve this for k in terms of j to find the ratio which must hold between $q^1{}_1$ and $q^1{}_2$ to enable introducing the price without changing the index. Algebraic reduction gives

$$\frac{k}{j}=\frac{ab(di-c)}{cd(a-bi)}=\frac{q^1{}_1}{q^1{}_2}.$$

We may therefore introduce a new commodity whose price ratio agrees with the index provided the quantities or weights are in the proportion shown. The following example is given as an illustration of the entirely reasonable weightings obtained. Here the $p^1{}_1q^1{}_1$ entered is large, being approximately one-tenth of the Σp_1q_1 term. Given

$$i=\sqrt{\frac{1800}{2000}\times\frac{2200}{2500}}=.889944$$

$$p^1{}_1=.20, \qquad q^1{}_1=1000$$

$$p^1{}_2=.2247333, \; q^1{}_2 \text{ is required.}$$

Solution gives $q^1{}_2=1236$, a very reasonable answer, as the ratio $\Sigma q_1/\Sigma q_2$ must very nearly equal 1000/1236, as it can hardly differ greatly from a/c or b/d, which equal $\frac{1000}{1222}$ and $\frac{1000}{1250}$ respectively.

The principle illustrated may be followed in introducing prices whose price ratio is not equal to the index. In such case k is not a rectilinear function of j, and if the introduced price ratio differs much from i quite absurd weightings might be required in order to preserve the value of the index. However, excepting only the weighted geometric mean index, this index lends itself the most readily to the introduction or withdrawal of items. The formula should be especially serviceable for chain indices.

$$i_{12}=\frac{\frac{\Sigma p_1q_1}{\Sigma p_2q_1}+\frac{\Sigma p_1q_2}{\Sigma p_2q_2}}{2}. \quad \text{(Index formula 14)}$$

By a procedure similar to the preceding for formula 15, it is found that a price ratio, equal to the index, may be introduced without changing the value of the index, if the weights or quantities $q^1{}_1$ and $q^1{}_2$ are in the ratio Σp_2q_1 to Σp_2q_2.

SCORES OF INDEX NUMBERS UPON BASIS OF SIX TESTS OF EXCELLENCE

Tests	(1) Type IA $\frac{1}{N}\Sigma\frac{p_1}{p_0}$ Carli Evelyn Economist Sauerbeck Soether	(2) Type IA $\frac{\Sigma\frac{p_1}{p_0}w}{\Sigma w}$ Young Falkner Dun	(3) Type IA $\frac{\Sigma\frac{p_1}{p_0}p_1q_1}{\Sigma p_1q_1}$ Palgrave	(4) Type II Median value of $\frac{p^1_1}{p^1_0}, \frac{p^2_1}{p^2_0} \dots$ Edgeworth	(5) Type II Weighted median	(6) Type III also Type V $\frac{\sqrt[n]{p^1_1 p^2_1 \dots}}{\sqrt[n]{p^1_2 n^2_2 \dots}}$ Jevons Westergaard	(7) Type III also Type V Weighted geom. mean	(8) Type IV $\frac{\Sigma p_1}{\Sigma p_2}$ Bradstreet
1 Reliability Smallness of P. E.	.5s-i	1.5s-i	1. s-i	2.	3.	1.	3.	1.5s-i
2 Parallels habitual mode of thinking	1.	1.	1.	1.	1.	.5	.5	2.
3 Proportionality	2.	2.	1.	2.	2.	2.	2.	2.
4 Entry and withdrawal	2.	2.	1.	1.–	1.–	2.	3.	2.
5 Change of base	.0	.0	.0	1.–	1.–	2.	2.	2.
6 Change of unit of measurement	2.	2.	2.	2.	2.	2.	2.	.0
Totals	7.5	8.5	6.0	9.0–	10.0–	9.5	12.5	9.5

SCORES OF INDEX NUMBERS UPON BASIS OF SIX TESTS OF EXCELLENCE—*Continued*

Tests	(9) Type IV $\frac{\Sigma p_1 w}{\Sigma p_2 w}$ Lowe	(10) Type IV $\frac{\Sigma p_1 \frac{q_0+q_1}{2}}{\Sigma p_0 \frac{q_0+q_1}{2}}$ Edgeworth Marshall	(11) Type IV $\frac{\Sigma p_1 \sqrt{q_0 q_1}}{\Sigma p_0 \sqrt{q_0 q_1}}$ Scrope and Walsh	(12) Type IV also Type IA $\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0}$ Scrope Sidgwick Sauerbeck Giffen	(13) Type IV also Type IH $\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1}$	(14) Type VI Arith. average of (12) and (13) Sidgwick Drobisch	(15) Type VI Geom. average of (12) and (13)	(16) Type V $\frac{\frac{\Sigma p_1 q_1}{\Sigma q_1}}{\frac{\Sigma p_2 q_2}{\Sigma q_2}}$ Drodisch Rawson-Rawson	(17) Type V $\frac{\frac{\Sigma p_1 q_1}{\sqrt[n]{q^1_1 q^1_1 \ldots}}}{\frac{\Sigma p_2 q_2}{\sqrt[n]{q^1_2 q^1_2 \ldots}}}$ Nicholson Walsh
1	2.5s-i	2.5s-i	2.5s-i	2. s-i	2. s-i	2.5s-i	2.5s-i	2. s-i	2. s-i
2	2.	1.5	1.5	1.	.5	.5	1.5	.5	.5
3	2.	1.	1.	2.	1.	1.	1.	.0	.0
4	2.	1.	1.	2.	1.	1.	1.	.0	.0
5	2.	1.	1.	.0	.0	.0	1.	2.	2.
6	2.–	2.	2.	2.	2.	2.	2.	.0	2.
Totals	12.5–	9.0	9.0	9.0	6.5	7.0	9.0	2.5	4.5

Type IA: Arithmetic average of ratios
Type IH: Harmonic average of ratios
Type II: Median of ratios
Type III: Geometric average
Type IV: Quotient of aggregates
Type V: Quotients of functions of data of single years
Type VI: Composites of preceding types

Test 5: Change of base. "The ratios between price indices should be unaffected by reversing or shifting the base." Algebraically stated: Let $\dot{i}_{12}=\frac{P_1}{P_2}$, $\dot{i}_{45}=\frac{P_4}{P_5}$, etc. Required that $\frac{\dot{i}_{34}}{\dot{i}_{14}}=\frac{\dot{i}_{32}}{\dot{i}_{12}}=\frac{P_3}{P_1}=\dot{i}_{31}$. Give score of 2 if true for any two years, score of 1 if only true when the base year and one other is involved, *i. e.*, if only such equations as $\frac{\dot{i}_{33}}{\dot{i}_{13}}=\dot{i}_{31}$, $\frac{\dot{i}_{22}}{\dot{i}_{42}}=\dot{i}_{24}$, etc., hold.

Test 6: Change of unit of measurement. "The ratios between various price indices should be unaffected by changing any unit of measurement." Score of 2 or 0.

Fisher has a "Determinateness" test which he describes in the words, "A price index should not be rendered zero, infinity, or indeterminate by an individual price becoming zero." This is but one phase of reliability and is therefore included in Test 1 above.

In the formulae listed the q's stand for quantities of commodities consumed or in trade and are weights of the p's. When weights not exactly equal to the q's are involved, the symbol w is used. It is of course assumed that care would be exercised in selecting these weights w. p_0 and q_0 instead of p_2 and q_2 are used in those formulae in which the treatment of the data for the base year is unique. Test 5 is not completely met by any such formulae.

Formulae 7 and 9, which are given the highest scores, involve weights, w, instead of quantities, q. There is great flexibility in each of these so that if a weight is adopted, let us say in the first instance upon the basis of quantities (if using formula 9) or values (if using formula 7) in trade, which tends to become unreasonable, it can be changed without affecting the index between the year when the change is made and the preceding year. If years from early to late are designated by 1, 2, 3, 4 and if a formula 7 index number is started at the end of the first year, using weights proportionate to the values of the commodities in trade, and continues until the beginning of year 4 before a change in weights is desirable, a change can at that time be made which will preserve the index $\dot{i}_{34}$ and its reciprocal $\dot{i}_{43}$. The new weighting would probably give an $\dot{i}_{42}$ and an $\dot{i}_{41}$, were they to be calculated, which would be slightly different from those given by the equations:

$$\dot{i}_{42}=\frac{\dot{i}_{43}}{\dot{i}_{23}} \text{ and } \dot{i}_{41}=\frac{\dot{i}_{43}}{\dot{i}_{13}}$$

which would exactly hold had no change in weights been made. This difference will usually be small, but if an index permitting changes in weightings and at the same time enabling the use, without approximation, of any year as base is demanded, it may be made by the expenditure of a little more labor.

Formula 12 (or 13) in which there are no parameters, or flexible weightings, will serve as a foundation:

$$i_{12} = \frac{\Sigma p_1 q_2}{\Sigma p_2 q_2}.$$

Let M_1 = the mean of the p_1's

m_1 = the mean of the q_1's

S_1 = the standard deviation of the p_1's

s_1 = the standard deviation of the q_1's

r_{11} = the correlation between the p_1's (represented by the first subscript) and the q_1's (represented by the second subscript).

Symbols with other subscripts have comparable meanings, *e. g.*, r_{24} = the correlation between the p_2's and the q_4's. Then,

$$\Sigma p_1 q_2 = N(M_1 m_2 + r_{12}\, S_1\, s_2)$$

$$\Sigma p_2 q_2 = N(M_2 m_2 + r_{22}\, S_2\, s_2).$$

Consequently, the numerator and the denominator for the index between any two years may be built up if the means, standard deviations, and correlations are known. The data required may be calculated each year, as the data for the years become available, and tabulated in such a table as the following:

DATA FOR DETERMINING INDEX WITH ANY DESIRED YEAR AS BASE

Years		Mp	Mq	$-S$	s	rpq: p for year indicated in stub and q for number of years indicated earlier (−) or later (+)												
						−32	−16	−8	−4	−2	−1	0	+1	+2	+4	+8	+16	+32
1919	9	+		+						+								
1918	8																	
1917	7	*	+×*	*	+×*							*						
1916	6																	
1915	5																	
1914	4																	
1913	3																	
1912	2																	
1911	1	×		×												×		

If it is desired to make 1917 the base and to express the prices in 1919 and 1911 relative to it, then $\Sigma p_9 q_7$ is determined from the magni-

tudes recorded in the compartments in which there is "+"; $\Sigma p_1 q_7$ from the compartments in which there is "$\times$"; and $\Sigma p_7 q_7$ from the compartments in which there is "*."

The table as drawn up does not provide space for all the possible correlation coefficients. With such care as could be taken in choosing the units of quantity, the correlation coefficients could be made to vary from year to year in a very regular manner, thus enabling interpolation with high accuracy. There is complete freedom in changing the weights of commodities, but it should be noted that a commodity "dropped" continues as one of zero price and zero quantity—in other words, the N has not been decreased by "dropping" the commodity. To change the weight of a commodity price from w to w' demands a warrant. Let us say that such warrant is found in the ratio of the quantities consumed. No less warrant is necessary when w' is zero. An article once included in the index should come out only in case it becomes practically obsolescent. No distortion of any index would result in this case. We may of course take out a commodity under other conditions without affecting some one particular index.

Formula 13, particularly serviceable if trade indices are involved, may be derived from the table of constants giving a formula 12 index.

The number and nature of the commodities entering into an index is a function of the accuracy required and of the particular purpose to be served by the index. Ruling out of consideration the index which is based upon a complete survey of a field the question is, what are the principles which should control in drawing a sampling? The fundamental principles of multiple correlation apply—high correlation with the purpose to be served and low intercorrelation. If a coal price index is being constructed from a small number drawn from a much larger number of quotations, the quotations should be chosen so that (a) each is as little correlated as possible with the other quotations included in the index, and (b) each is as highly correlated as possible with the other quotations in the field not included in the index. It is to be expected that commercial tendencies will conspire to prevent any quotation from markedly possessing both characteristics, in which case a balance must be struck between them: (b) is the more important if the number of quotations in the index is mall, say not over six, but (a) is by far the more important if the number of quotations is large. In fact, quotations that are excellent for incorporation in an index number based upon a small number of items may be expected to be relatively inferior for incorporation in an index based upon a large number of items. This brief observation as to the significance of correlation between commodity prices is, in the main, an addendum to,

not in opposition to, the points involved in Dr. Mitchell's (1915) very thorough exposition of the question of what commodities should be included.

GENERAL INDEX NUMBERS

Brown, William: *Mental Measurement.* Cambridge University Press. 1911.

Edgeworth, F. Y.: "Index Numbers," *Dictionary of Political Economy.* v. 2.

—— ——: Memoranda subjoined to *Reports of the British Association,* 1887, pp. 247–54; 1888, pp. 181–8; 1889, p. 133; 1890, pp. 485–8.

Fisher, Irving: *Purchasing Power of Money.* Revised ed. Macmillan Co. 1913. (1st ed. 1911.)

—— ——: "Best Form of Index Number." *Qu. Pub. Am. Stat. Assn.* March, 1921.

Kelley, Truman L.: "A Simplified Method of Using Scaled Data for Purposes of Testing." *Sch. and Soc.* vol. 4, nos. 79–80, 1916.

Knibbs, G. H.: *Prices, Price Indexes and Cost of Living in Australia.* Bur. of Census and Stat., Labor and Indust. Br. Report No. 1, 1912. Also Report No. 9, 1918.

Mitchell, Wesley C.: Author of Part 1.—*The Making and Use of Index Numbers.* Bul. U. S. Bureau of Labor Stat., whole number 173. 1915.

Pearson, Karl: "On the Constants of Index-distribution as deduced from the like constants for the Components of the Ratio." *Biometrika,* vol. 7, 1910.

—— ——: "The Opsonic Index." *Biometrika,* vol. 8, 1911.

Walsh, C. M.: *Measurement of General Exchange Value.* Macmillan. 1910.

—— ——: *Problem of Estimation.* London, P. S. King, 1921.

Yule, G. Udney: *Introduction to the Theory of Statistics.* Lippincott. 1912.